Yoga Sundaram Rama Swamy

216-200-2712 | yogasundaram95@gmail.com | linkedin.com/in/yogasundaram | github.com/yogasundaram95

Open to relocation nationwide | Available for on-site, hybrid, or remote roles

Data Scientist | Data Analyst

Data Analyst with 5+ years of experience analyzing customer, CRM, and behavioral data to drive campaign performance and KPI outcomes. Skilled in Python, R, SQL, and predictive modeling with strong foundation in hypothesis testing, experimental design, and translating analytical findings into decision-ready insights for business and senior leadership.

Experience

Sales Operations Data Analyst & Business Intelligence Co-op
Jan 2025 – Present

Bendix Commercial Vehicle Systems — *Avon, OH*

- Built executive-level Power BI dashboards with standardized KPI definitions and DAX measures, ensuring **99%+ data accuracy** for leadership reporting and performance tracking
- Analyzed campaign and engagement data using the WebTrack application to monitor page views, user interactions, and conversion-related events, enabling KPI tracking and optimization insights that improved campaign visibility and accelerated stakeholder decision-making
- Developed workload and effort estimation views using cycle-time, throughput, and interaction frequency to support exploratory capacity planning and staffing discussions across Sales Operations
- Improved reporting cadence alignment with demand cycles, reducing unplanned backlog visibility delays by **18%** (internal KPI).
- Built forecasting prototypes in Power BI and Python (ARIMA/ETS) to anticipate usage and demand trends; enabled data-driven prioritization for routing, resource allocation, and reporting cadence
- Partnered with cross-functional stakeholders to define workload categories and preliminary duration assumptions, laying groundwork for time-to-event modeling and staffing optimization initiatives
- Enhanced ETL pipelines for large-scale marketing asset migration (**100K+ records**) by adding schema validation, completeness checks, and drift monitoring to ensure reproducibility and decision-grade data integrity
- Automated survey and feedback workflows and created KPI-quality gates, reducing manual collection by **2+ hours/day** and improving input reliability for operational analysis

Machine Learning Data Analyst
Oct 2020 – Mar 2023

Amazon Development Centre — *Chennai, TN*

- Extracted and transformed large-scale behavioral datasets using SQL (Amazon Redshift) and Python to support KPI tracking and ML performance analysis
- Applied regression modeling, GLM concepts (logistic/Poisson), and time-to-event reasoning to analyze workflow duration patterns, predict completion likelihood, and identify high-impact bottlenecks in annotation cycles
- Designed real-time operational dashboards (QuickSight) to monitor queue lengths, throughput, and failure rates, enabling scenario-based staffing conversations and informing assignment strategy experiments
- Audited analytical code and workflows to ensure data integrity, consistency, and compliance with quality standards
- Conducted A/B testing on workflow and model changes using control vs. treatment groups across **10K+ samples**, applying hypothesis testing to confirm **10% throughput improvement** and measurable gains in ML training data quality
- Trained and mentored **20+ associates**, improving annotation throughput by **10%** and standardizing onboarding processes

Data Analyst
Jul 2017 – Oct 2019

Gharib Shipping Services LLC — *Chennai, TN*

- Designed and implemented comprehensive data collection processes for warehouse operations, analyzing 70,000+ rows of historical shipping and inventory data using SQL to identify performance trends and operational bottlenecks, resulting in 15% efficiency improvement and $2K monthly cost reduction.
- Established data protocols for shipping and inventory data collection, ensuring data quality and consistency across multiple operational systems while supporting strategic decision-making through comprehensive business intelligence solutions.

Education

Master of Information Systems
Aug 2023 – Dec 2025

Cleveland State University — *Cleveland, OH*

GPA: 3.94/4.00

Bachelor of Technology in Aeronautical Engineering
Jun 2013 – Jun 2017

BSA Crescent Institute of Science and Technology — *Chennai, India*

Selected Analytics & Research Projects

Time-Series Stock Prediction Using Ensemble ML Model | *Python, XGBoost, LSTM, ARIMA, Scikit-learn* 2024 − 2025
- Led independent research integrating 5+ heterogeneous data sources (financial markets, macroeconomic indicators, sentiment analysis); designed experimental framework with ensemble ML methods (Random Forest, XGBoost, LSTM) and conducted SHAP feature importance analysis with rigorous time-series cross-validation on 5+ years of daily observations.
- Achieved **94% directional accuracy ($R^2 = 0.88$)** through systematic feature engineering and model optimization; validated external factors improved predictive performance by 8-12% through ablation studies; documented methodology and findings in comprehensive technical report with reproducible code.

Statistical Modeling for Credit Risk Classification | *Weka, R, SQL, Predictive Analytics* **Jan 2025 − Feb 2025**
- Designed and executed comparative study of classification algorithms (Decision Tree, Neural Networks) for credit risk assessment; performed systematic hyperparameter optimization and preprocessing to improve model performance from baseline 64-66% to 68-70% accuracy.
- Conducted rigorous model validation using 10-fold cross-validation and Leave-One-Out methods; performed ROC analysis for threshold optimization and implemented cost-sensitive learning frameworks; achieved 80-100% accuracy on held-out test set with AUC = 0.814.

Relevant Courses

- Data Visualization Theory and Practice
- Business Intelligence Systems
- Enterprise Databases
- Machine Learning
- Modern Database Design
- Advanced Systems Analysis
- Enterprise Resource Planning
- Project Management

Technical Skills

Analytics & BI: Power BI (DAX, Power Query), Tableau, QuickSight, KPI dashboards, operational/campaign performance analysis
Programming: SQL (advanced queries, Redshift), Python (pandas, NumPy, scikit-learn, statsmodels), R, SAS
Data Platforms: AWS Redshift, SAP HANA, Snowflake, ETL pipelines, data quality & validation, reproducible workflows
Advanced Analytics: Statistical modeling, hypothesis testing, experimentation, feature engineering, SHAP
Modeling & OR: GLM (logistic/Poisson), time-to-event concepts (Cox/KM — learning), queueing fundamentals, ARIMA/ETS forecasting, discrete-event simulation exposure (SimPy — exploratory)

Soft Skills

attention to detail & accuracy; analytical problem-solving; hypothesis-driven investigation; clear written & verbal communication; technical documentation & research reporting; organization & time management; stakeholder collaboration & presenting analytical findings; proactive issue escalation; meticulous documentation

Certifications

Google Data Analytics Professional Certificate	*Oct 2024*
Microsoft Certified: Fabric Data Engineer Associate	*Dec 2025*
Oracle Cloud Data Management Certified Foundations Associate	*Nov 2023*

Additional Information

In OPT including STEM extension (valid through Dec 2029). Willing to relocate nationwide. Authorized to work with any employer in the U.S.